

Patent Application US/07/599,543B

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Opperman, Hermann
Ozkaynak, Engin
Rueger, David C.
Kuberasampath, Thangavel
- (ii) TITLE OF INVENTION: Osteogenic Proteins
- (iii) NUMBER OF SEQUENCES: 11
- (iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Testa Hurwitz & Thibault
(B) STREET: 53 State Street
(C) CITY: Boston
(D) STATE: Massachusetts
(E) COUNTRY: U.S.A.
(F) ZIP: 02109
- (v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
(B) COMPUTER: IBM XT
(C) OPERATING SYSTEM: DOS 3.30
(D) SOFTWARE: ASC II
- (vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: US 07/599,543
(B) FILING DATE: 18-Oct-90
- (vii) ~~PRIOR APPLICATION DATA:~~ *insert*
(A) APPLICATION NUMBER: US 569,920
(B) FILING DATE: 20-Aug-90
(C) APPLICATION NUMBER: US 315,342
(D) FILING DATE: 23-Feb-89
(E) APPLICATION NUMBER: US 422,699
(F) FILING DATE: 17-Oct-89

**Does not comply
- Send a notice -****(C) CLASSIFICATION:**

Raw Sequence Listing
Patent Application US/07/599,543B07/08/91
11:17:57

54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71 (2) INFORMATION FOR SEQ ID NO:1:
72 (i) SEQUENCE CHARACTERISTICS:
73 (A) LENGTH: 139 amino acids
74 (B) TYPE: amino acid
75 (D) TOPOLOGY: linear
76 (ii) MOLECULE TYPE: protein
77 (ix) FEATURE:
78 (A) NAME: mOP2 (mature)
79 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
80
81 Ala Ala Arg Pro Leu Lys Arg Arg Gln
82 1 5
83 Pro Lys Lys Thr Asn Glu Leu Pro His
84 10 15
85 Pro Asn Lys Leu Pro Gly Ile Phe Asp
86 20 25
87 Asp Gly His Gly Ser Arg Gly Arg Glu
88 30 35
89 Val Cys Arg Arg His Glu Leu Tyr Val
90 40 45
91 Arg Phe Arg Asp Leu Gly Trp Leu Asp
92 50
93 Trp Val Ile Ala Pro Gln Gly Tyr Ser
94 55 60
95 Ala Tyr Tyr Cys Glu Gly Glu Cys Ala
96 65 70
97 Phe Pro Leu Asp Ser Cys Met Asn Ala
98 75 80
99 Thr Asn His Ala Ile Leu Gln Ser Leu
100 85 90
101 Val His Leu Met Lys Pro Asp Val Val
102 95
103 Pro Lys Ala Cys Cys Ala Pro Thr Lys
104 100 105
105 Leu Ser Ala Thr Ser Val Leu Tyr Tyr
106 110 115

Raw Sequence Listing
Patent Application US/07/599,543B

07/08/91
11:17:59

107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136

137	Asp	Ser	Ser	Asn	Asn	Val	Ile	Leu	Arg
138			120					125	
139	Lys	His	Arg	Asn	Met	Val	Val	Lys	Ala
140				130					135
141	Cys	Gly	Cys	His					

142
143

144 (2) INFORMATION FOR SEQ ID NO:2:
145 (i) SEQUENCE CHARACTERISTICS:
146 (A) LENGTH: 1930 base pairs
147 (B) TYPE: nucleic acid
148 (C) STRANDEDNESS: single
149 (D) TOPOLOGY: linear
150 (ii) MOLECULE TYPE: cDNA
151 (iii) HYPOTHETICAL: no
152 (iv) ANTI-SENSE: no
153 (vi) ORIGINAL SOURCE:
154 (A) ORGANISM: mouse
155 (F) TISSUE TYPE: embryo
156 (ix) FEATURE:
157 (A) NAME: mOP2
158 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:
159

Patent Application US/07/599,543B

```
160 GGAATTCGCG TGCCAGGCAC AGGTGCGCCG TCTGGTCCTC 40
161 CCCGTCTGGC GTCAGCCGAG CCCGACCAGC TACCAGTGGA 80
162 TGCGCGCCGG CTGAAAGTCC GAG ATG GCT ATG CGT 115
163 Met Ala Met Arg
164 1
165 CCC GGG CCA CTC TGG CTA TTG GGC CTT GCT CTG 148
166 Pro Gly Pro Leu Trp Leu Leu Gly Leu Ala Leu
167 5 10 15
168 TGC GCG CTG GGA GGC GGC CAC GGT CCC GGT CCC 181
169 Cys Ala Leu Gly Gly Gly His Gly Pro Gly Pro
170 20 25
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202 CCG CAC ACC TGT CCC CAG CGT CGC CTG GGA GCG 214
203 Pro His Thr Cys Pro Gln Arg Arg Leu Gly Ala
204 30 35
205 CGC GAC CGG GAC ATG CAG CGT GAA ATC CTG CCG 247
206 Arg Asp Arg Asp Met Gln Arg Glu Ile Leu Pro
207 40 45
208 GTG CTC GGG CTA CCG GGA CGC CCC GAC CCC GTG 280
209 Val Leu Gly Leu Pro Gly Arg Pro Asp Pro Val
210 50 55
211 CAC AAC CCG CCG CTG CCC GGC ACG CAG CGT GCG 313
212 His Asn Pro Pro Leu Pro Gly Thr Gln Arg Ala
```

Patent Application US/07/599,543B

[illegible]

Patent Application US/07/599,543B

266
267 AGC ACC CAC CCG CTC AAC ACA ACC CTC CAC ATC 577
268 Ser Thr His Pro Leu Asn Thr Thr Leu His Ile
269 150 155
270 AGC ATG TTC GAA GTG GTC CAA GAG CAC TCC AAC 610
271 Ser Met Phe Glu Val Val Gln Glu His Ser Asn
272 160 165
273 AGG GAG TCT GAC TTG TTC TTT TTG GAT CTT CAG 643
274 Arg Glu Ser Asp Leu Phe Phe Leu Asp Leu Gln
275 170 175 180
276 ACG CTC CGA TCT GGG GAC GAG GGC TGG CTG GTG 676
277 Thr Leu Arg Ser Gly Asp Glu Gly Trp Leu Val
278 185 190
279 CTG GAC ATC ACA GCA GCC AGT GAC CGA TGG CTG 709
280 Leu Asp Ile Thr Ala Ala Ser Asp Arg Trp Leu
281 195 200
282 CTG AAC CAT CAC AAG GAC CTG GGA CTC CGC CTC 742
283 Leu Asn His His Lys Asp Leu Gly Leu Arg Leu
284 205 210
285 TAT GTG GAA ACC GCG GAT GGG CAC AGC ATG GAT 775
286 Tyr Val Glu Thr Ala Asp Gly His Ser Met Asp
287 215 220
288 CCT GGC CTG GCT GGT CTG CTT GGA CGA CAA GCA 808
289 Pro Gly Leu Ala Gly Leu Leu Gly Arg Gln Ala
290 225 230 235
291 CCA CGC TCC AGA CAG CCT TTC ATG GTA ACC TTC 841
292 Pro Arg Ser Arg Gln Pro Phe Met Val Thr Phe
293 240 245
294 TTC AGG GCC AGC CAG AGT CCT GTG CGG GCC CCT 874
295 Phe Arg Ala Ser Gln Ser Pro Val Arg Ala Pro
296 250 255
297 CGG GCA GCG AGA CCA CTG AAG AGG AGG CAG CCA 907
298 Arg Ala Ala Arg Pro Leu Lys Arg Arg Gln Pro
299 260 265
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318

Patent Application US/07/599,543B

319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332 AAG AAA ACG AAC GAG CTT CCG CAC CCC AAC AAA 940
 333 Lys Lys Thr Asn Glu Leu Pro His Pro Asn Lys
 334 270 275
 335 CTC CCA GGG ATC TTT GAT GAT GGC CAC GGT TCC 973
 336 Leu Pro Gly Ile Phe Asp Asp Gly His Gly Ser
 337 280 285 290
 338 CGC GGC AGA GAG GTT TGC CGC AGG CAT GAG CTC 1006
 339 Arg Gly Arg Glu Val Cys Arg Arg His Glu Leu
 340 295 300
 341 TAC GTC AGA TTC CGT GAC CTT GGC TGG CTG GAC 1039
 342 Tyr Val Arg Phe Arg Asp Leu Gly Trp Leu Asp
 343 305 310
 344 TGG GTC ATC GCC CCC CAG GGC TAC TCT GCC TAT 1072
 345 Trp Val Ile Ala Pro Gln Gly Tyr Ser Ala Tyr
 346 315 320
 347 TAC TGT GAG GGG GAG TGT GCT TTC CCA CTG GAC 1105
 348 Tyr Cys Glu Gly Glu Cys Ala Phe Pro Leu Asp
 349 325 330
 350 TCC TGT ATG AAC GCC ACC AAC CAT GCC ATC TTG 1138
 351 Ser Cys Met Asn Ala Thr Asn His Ala Ile Leu
 352 335 340 345
 353 CAG TCT CTG GTG CAC CTG ATG AAG CCA GAT GTT 1171
 354 Gln Ser Leu Val His Leu Met Lys Pro Asp Val
 355 350 355
 356 GTC CCC AAG GCA TGC TGT GCA CCC ACC AAA CTG 1204
 357 Val Pro Lys Ala Cys Cys Ala Pro Thr Lys Leu
 358 360 365
 359 AGT GCC ACC TCT GTG CTG TAC TAT GAC AGC AGC 1237
 360 Ser Ala Thr Ser Val Leu Tyr Tyr Asp Ser Ser
 361 370 375
 362 AAC AAT GTC ATC CTG CGT AAA CAC CGT AAC ATG 1270
 363 Asn Asn Val Ile Leu Arg Lys His Arg Asn Met
 364 380 385
 365
 366
 367
 368
 369
 370
 371

Patent Application US/07/599,543B

372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397 GTG GTC AAG GCC TGT GGC TGC CAC 1294
398 Val Val Lys Ala Cys Gly Cys His
399 390 400
400 TGAGGCCCG CCCAGCATCC TGCTTCTACT ACCTTACCAT 1334
401 CTGGCCGGGC CCCTCTCCAG AGGCAGAAAC CCTTCTATGT 1374
402 TATCATAGCT CAGACAGGGG CAATGGGAGG CCCTTCACTT 1414
403 CCCCTGGCCA CTTCTGCTA AAATTCTGGT CTTTCCCAGT 1454
404 TCCTCTGTCC TTCATGGGGT TTCGGGGCTA TCACCCCGCC 1494
405 CTCTCCATCC TCCTACCCCA AGCATAGACT GAATGCACAC 1534
406 AGCATCCAG AGCTATGCTA ACTGAGAGGT CTGGGGTCAG 1574
407 CACTGAAGGC CCACATGAGG AAGACTGATC CTTGGCCATC 1614
408 CTCAGCCCAC AATGGCAAAT TCTGGATGGT CTAAGAAGCC 1654
409 CTGGAATTCT AAAC TAGATG ATCTGGGCTC TCTGCACCAT 1694
410 TCATTGTGGC AGTTGGGACA TTTT TAGGTA TAACAGACAC 1734
411 ATACACTTAG ATCAATGCAT CGCTGTACTC CTTGAAATCA 1774
412 GAGCTAGCTT GTTAGAAAA GAATCAGAGC CAGGTATAGC 1814
413 GGTGCATGTC ATTAATCCCA GCGCTAAAGA GACAGAGACA 1854
414 GGAGAATCTC TGTGAGTTCA AGGCCACATA GAAAGAGCCT 1894
415 GTCTCGGGAG CAGGAAAAA AAAAAAACG GAATTC 1930
416
417

418 (2) INFORMATION FOR SEQ ID NO:3:
419 (i) SEQUENCE CHARACTERISTICS:
420 (A) LENGTH: 139 amino acids
421 (B) TYPE: amino acid
422 (D) TOPOLOGY: linear
423 (ii) MOLECULE TYPE: protein
424 (ix) FEATURE:

Patent Application US/07/599,543B

425 (A) NAME: hOP2 (mature)
426 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:
427
428 Ala Val Arg Pro Leu Arg Arg Arg Gln
429 1 5
430 Pro Lys Lys Ser Asn Glu Leu Pro Gln
431 10 15
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464 Ala Asn Arg Leu Pro Gly Ile Phe Asp
465 20 25
466 Asp Val His Gly Ser His Gly Arg Gln
467 30 35
468 Val Cys Arg Arg His Glu Leu Tyr Val
469 40 45
470 Ser Phe Gln Asp Leu Gly Trp Leu Asp
471 50
472 Trp Val Ile Ala Pro Gln Gly Tyr Ser
473 55 60
474 Ala Tyr Tyr Cys Glu Gly Glu Cys Ser
475 65 70
476 Phe Pro Leu Asp Ser Cys Met Asn Ala
477 75 80

Patent Application US/07/599,543B

478 Thr Asn His Ala Ile Leu Gln Ser Leu
479 85 90
480 Val His Leu Met Lys Pro Asn Ala Val
481 95
482 Pro Lys Ala Cys Cys Ala Pro Thr Lys
483 100 105
484 Leu Ser Ala Thr Ser Val Leu Tyr Tyr
485 110 115
486 Asp Ser Ser Asn Asn Val Ile Leu Arg
487 120 125
488 Lys Ala Arg Asn Met Val Val Lys Ala
489 130 135
490 Cys Gly Cys His
491
492

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1956 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

only 1941 are listed
see line 797

493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

Patent Application US/07/599,543B

531 (iii) HYPOTHETICAL: no
532 (iv) ANTI-SENSE: no
533 (vi) ORIGINAL SOURCE:
534 (A) ORGANISM: homo sapiens
535 (F) TISSUE TYPE: hippocampus
536 (ix) FEATURE:
537 (A) NAME: hOP2
538 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:
539
540 GGAATTCCGG CCACAGTGGC GCCGGCAGAG CAGGAGTGGC 40
541 TGGAGGAGCT GTGGTTGGAG CAGGAGGTGG CACGGCAGGG 80
542 CTGGAGGGCT CCCTATGAGT GCGGGAGACG GCCCAGGAGG 120
543 CGCTGGAGCA ACAGCTCCCA CACCGCACCA AGCGGTGGCT 160
544 GCAGGAGCTC GCCCATCGCC CCTGCGCTGC TCGGACCGCG 200
545 GCCACAGCCG GACTGGCGGG TACGGCGGCG ACAGAGGCAT 240
546 TGGCCGAGAG TCCCAGTCCG CAGAGTAGCC CCGGCCTCGA 280
547 GCGGGTGGCG TCCCGGTCCT CTCCGTCCAG GAGCCAGGAC 320
548 AGGTGTCGCG CGGCGGGGCT CCAGGGACCG CGCCTGAGGC 360
549 CGGCTGCCCC CCGTCCCCG CCGCCCCCG CGCCCCCGCG 400
550 CCGCCGAGCC CAGCCTCCTT GCCGTGCGGG CGTCCCCAGG 440
551 CCCTGGGTCG GCCGCGGAGC CGATGCGCGC CCGCTGAGCG 480
552 CCCCAGCTGA GCGCCCCCGG CCTGCC ATG ACC GCG CTC 518
553 Met Thr Ala Leu
554 1
555 CCC GGC CCG CTC TGG CTC CTG GGC CTG GCG CTA 551
556 Pro Gly Pro Leu Trp Leu Leu Gly Leu Ala Leu
557 5 10 15
558 TGC GCG CTG GGC GGG GGC GGC CCC GGC CTG CGA 584
559 Cys Ala Leu Gly Gly Gly Gly Pro Gly Leu Arg
560 20 25
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583

Patent Application US/07/599,543B

```

584
585
586
587
588
589
590
591
592
593
594   CCC CCG CCC GGC TGT CCC CAG CGA CGT CTG GGC   617
595   Pro Pro Pro Gly Cys Pro Gly arg Arg Leu Gly
596           30                               35
597   GCG CGC GAC CGG GAC GTG CAG CGC GAG ATC CTG   650
598   Ala Arg Asp Arg Asp Val Gln Arg Glu Ile Leu
599           40                               45
600   GCG GTG CTC GGG CTG CCT GGG CGG CCC CGG CCC   683
601   Ala Val Leu Gly Leu Pro Gly Arg Pro Arg Pro
602           50                               55
603   CGC GCG CCA CCC GCC GCC TCC CGG CTG CCC GCG   716
604   Arg Ala Pro Pro Ala Ala Ser Arg Leu Pro Ala
605           60                               65                               70
606   TCC GCG CCG CTC TTC ATG CTG GAC CTG TAC CAC   749
607   Ser Ala Pro Leu Phe Met Leu Asp Leu Tyr His
608           75                               80
609   CGC ATG GCC GGC GAC GAC GAC GAG GAC GGC GCC   782
610   Arg Met Ala Gly Asp Asp Asp Glu Asp Gly Ala
611           85                               90
612   GCG GAG GCC CTG GGC CGC GCC GAC CTG GTC ATG   815
613   Ala Glu Ala Leu Gly Arg Ala Asp Leu Val Met
614           95                               100
615   AGC TTC GTT AAC ATG GTG GAG CGA GAC CGT GCC   848
616   Ser Phe Val Asn Met Val Glu Arg Asp Arg Ala
617           105                              110
618   CTG GGC CAC CAG GAG CCC CAT TGG AAG GAG TTC   881
619   Leu Gly His Gln Glu Pro His Trp Lys Glu Phe
620           115                              120                              125
621   CGC TTT GAC CTG ACC CAG ATC CCG GCT GGG GAG   914
622   Arg Phe Asp Leu Thr Gln Ile Pro Ala Gly Glu
623           130                              135
624   GCG GTC ACA GCT GCG GAG TTC CGG ATT TAC AAG   947
625   Ala Val Thr Ala Ala Glu Phe Arg Ile Tyr Lys
626           140                              145
627
628
629
630
631
632
633
634
635
636

```

Patent Application US/07/599,543B

637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659 GTG CCC AGC ATC CAC CTG CTC AAC AGG ACC CTC 980
660 Val Pro Ser Ile His Leu Leu Asn Arg Thr Leu
661 150 155
662 CAC GTC AGC ATG TTC CAG GTG GTC CAG GAG CAG 1013
663 His Val Ser Met Phe Gln Val Val Gln Glu Gln
664 160 165
665 TCC AAC AGG GAG TCT GAC TTG TTC TTT TTG GAT 1046
666 Ser Asn Arg Glu Ser Asp Leu Phe Phe Leu Asp
667 170 175 180
668 CTT CAG ACG CTC CGA GCT GGA GAC GAG GGC TGG 1079
669 Leu Gln Thr Leu Arg Ala Gly Asp Glu Gly Trp
670 185 190
671 CTG GTG CTG GAT GTC ACA GCA GCC AGT GAC TGC 1112
672 Leu Val Leu Asp Val Thr Ala Ala Ser Asp Cys
673 195 200
674 TGG TTG CTG AAG CGT CAC AAG GAC CTG GGA CTC 1145
675 Trp Leu Leu Lys Arg His Lys Asp Leu Gly Leu
676 205 210
677 CGC CTC TAT GTG GAG ACT GAG GAC GGG CAC AGC 1178
678 Arg Leu Tyr Val Glu Thr Glu Asp Gly His Ser
679 215 220
680 GTG GAT CCT GGC CTG GCC GGC CTG CTG GGT CAA 1211
681 Val Asp Pro Gly Leu Ala Gly Leu Leu Gly Gln
682 225 230 235
683 CGG GCC CCA CGC TCC CAA CAG CCT TTC GTG GTC 1244
684 Arg Ala Pro Arg Ser Gln Gln Pro Phe Val Val
685 240 245
686 ACT TTC TTC AGG GCC AGT CCG AGT CCC ATC CGC 1277
687 Thr Phe Phe Arg Ala Ser Pro Ser Pro Ile Arg
688 250 255
689 ACC CCT CGG GCA GTG AGG CCA CTG AGG AGG AGG 1310

Patent Application US/07/599,543B

690	Thr	Pro	Arg	Ala	Val	Arg	Pro	Leu	Arg	Arg	Arg	
691			260					265				
692												
693												
694												
695												
696												
697												
698												
699												
700												
701												
702												
703												
704												
705												
706												
707												
708												
709												
710												
711												
712												
713												
714												
715												
716												
717												
718												
719												
720												
721												
722												
723												
724	CAG	CCG	AAG	AAA	AGC	AAC	GAG	CTG	CCG	CAG	GCC	1343
725	Gln	Pro	Lys	Lys	Ser	Asn	Glu	Leu	Pro	Gln	Ala	
726			270					275				
727	AAC	CGA	CTC	CCA	GGG	ATC	TTT	GAT	GAC	GTC	CAC	1376
728	Asn	Arg	Leu	Pro	Gly	Ile	Phe	Asp	Asp	Val	His	
729			280				285				290	
730	GGC	TCC	CAC	GGC	CGG	CAG	GTC	TGC	CGT	CGG	CAC	1409
731	Gly	Ser	His	Gly	Arg	Gln	Val	Cys	Arg	Arg	His	
732					295					300		
733	GAG	CTC	TAC	GTC	AGC	TTC	CAG	GAC	CTC	GGC	TGG	1442
734	Glu	Leu	Tyr	Val	Ser	Phe	Gln	Asp	Leu	Gly	Trp	
735				305						310		
736	CTG	GAC	TGG	GTC	ATC	GCT	CCC	CAA	GGC	TAC	TCG	1475
737	Leu	Asp	Trp	Val	Ile	Ala	Pro	Gln	Gly	Tyr	Ser	
738			315					320				
739	GCC	TAT	TAC	TGT	GAG	GGG	GAG	TGC	TCC	TTC	CCA	1508
740	Ala	Tyr	Tyr	Cys	Glu	Gly	Glu	Cys	Ser	Phe	Pro	
741			325				330					
742	CTG	GAC	TCC	TGC	ATG	AAT	GCC	ACC	AAC	CAC	GCC	1541

Patent Application US/07/599,543B

743 Leu Asp Ser Cys Met Asn Ala Thr Asn His Ala
744 335 340 345
745 ATC CTG CAG TCC CTG GTG CAC CTG ATG AAG CCA 1574
746 Ile Leu Gln Ser Leu Val His Leu Met Lys Pro
747 350 355
748 AAC GCA GTC CCC AAG GCG TGC TGT GCA CCC ACC 1607
749 Asn Ala Val Pro Lys Ala Cys Cys Ala Pro Thr
750 360 365
751 AAG CTG AGC GCC ACC TCT GTG CTC TAC TAT GAC 1640
752 Lys Leu Ser Ala Thr Ser Val Leu Tyr Tyr Asp
753 370 375
754 AGC AGC AAC AAC GTC ATC CTG CGC AAA GCC CGC 1673
755 Ser Ser Asn Asn Val Ile Leu Arg Lys Ala Arg
756 380 385
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789 AAC ATG GTG GTC AAG GCC TGC GGC TGC CAC 1703
790 Asn Met Val Val Lys Ala Cys Gly Cys His
791 390 395
792 TGAGTCAGCC CGCCCAGCCC TACTGCAGCA ATTCACTGGC 1743
793 CGTCGTTTTA CAACGTGTGA CTGGGAAAAC CCTGGCGTTA 1783
794 CCCAACTTAA TCGCCTTGCA GCACATCCCC CTTTCGCCAG 1823
795 CTGGCTAATA GCGAAGAGGC CCCGCACCGA TCGCCCTTCC 1863

Patent Application US/07/599,543B

796 CAACAGTTGC GCCCCAGTGA ATGGCGAATG GCAAATTGTA 1903
797 AGCGTTAATA TTTTGTTAAA ATTCGCGTTA AATTTTTT 1941

see line 495

800 (2) INFORMATION FOR SEQ ID NO:5:

801 (i) SEQUENCE CHARACTERISTICS:

802 (A) LENGTH: 98 amino acids

803 (B) TYPE: amino acid

804 (D) TOPOLOGY: linear

805 (ii) MOLECULE TYPE: protein

806 (ix) FEATURE:

807 (D) OTHER INFORMATION: wherein "res."

808 means "residue" and Xaa at res. 2 = (Lys or Arg); Xaa at
809 res.3 = (Lys or Arg); Xaa res.9 = (Ser or Arg); Xaa at
810 res.11 = (Arg or Gln); Xaa at res.16 = (Gln or Leu); Xaa
811 at res. 19 = (Ile or Val); Xaa at res.23 = (Glu or Gln);
812 Xaa at res.26 = (Ala or Ser); Xaa at res. 34 = (Ala or
813 or Ser); Xaa at res.38= (Asn or Asp); Xaa at res. 40 =
814 (Tyr or Cys); Xaa at res.49 = (Val or Leu); Xaa at
815 res.52= (His or Asn); Xaa at res. 53 = (Phe or
816 Leu); Xaa at res. 54 = (Ile or Met); Xaa at res. 55 = (Asn
817 or Lys); Xaa at res. 56 = (Glu, Asp or Asn); Xaa at res.
818 57=(Thr, Ala or Val); Xaa at res. 61 = (Pro or Ala);
819 Xaa at res. 67=(gln or Lys); Xaa at res. 69 =
820 (Asn or Ser); Xaa at 71=(Ile or Thr); Xaa at res.
821 76= (Phe or Tyr); Xaa at res. 78 = (Asp, Glu or Ser);
822 Xaa at res. 80= (Ser or Asn); Xaa at res. 84 = (Ile or
823 Asp); Xaa at res. 85 Arg); Xaa at res. 87 = (Tyr, Ala
824 or His); and Xaa at res. 93=(Arg or Lys)

OK
8/2/91
7/2/91

826 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

827
828 Cys Xaa Xaa His Glu Leu Tyr Val Xaa Phe
829 1 5 10
830 Xaa Asp Leu Gly Trp Xaa Asp Trp Xaa Ile
831 15 20
832 Ala Pro Xaa Gly Tyr Xaa Ala Tyr Tyr Cys
833 25 30
834 Glu Gly Cys Xaa Phe Pro Leu Xaa Ser Xaa
835 35 40
836 Met Asn Ala Thr Asn His Ala Ile Xaa Thr
837 45 50
838 Leu Xaa Xaa Xaa Xaa Xaa Xaa Val
839 55
840 Pro Lys Xaa Cys Cys Ala Pro Thr Xaa Leu
841 60 65

Raw Sequence Listing
Patent Application US/07/599,543B

07/08/91
11:19:26

849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901

Xaa	Ala	Xaa	Ser	Val	Leu	Tyr	Xaa	Asp
	70				75			
Xaa	Ser	Xaa	Asn	Val	Xaa	Leu	Xaa	Lys
	80				85			
Xaa	Pro	Asn	Met	Val	Val	Xaa	Ala	Cys Gly
		90				95		
Cys	His							

- (2) INFORMATION FOR SEQ ID NO:6:
- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 437 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: cDNA
 - (iii) HYPOTHETICAL: no
 - (iv) ANTI-SENSE: no
 - (vi) ORIGINAL SOURCE:

Patent Application US/07/599,543B

902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921 (A) ORGANISM: Human
922 (F) TISSUE TYPE: placenta
923 (ix) FEATURE:
924 (A) NAME: OP1
925 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:
926
927 TCC ACG GGG 9
928 Ser Thr Gly
929 1
930 AGC AAA CAG CGC AGC CAG AAC CGC TCC AAG ACG 42
931 Ser Lys Gln Arg Ser Gln Asn Arg Ser Lys Thr
932 5 10
933 CCC AAG AAC CAG GAA GCC CTG CGG ATG GCC AAC 75
934 Pro Lys Asn Gln Glu Ala Leu Arg Met Ala Asn
935 15 20 25
936 GTG GCA GAG AAC AGC AGC AGC GAC CAG AGG CAG 108
937 Val Ala Glu Asn Ser Ser Ser Asp Gln Arg Gln
938 30 35
939 GCC TGT AAG AAG CAC GAG CTG TAT GTC AGC TTC 141
940 Ala Cys Lys Lys His Glu Leu Tyr Val Ser Phe
941 40 45
942 CGA GAC CTG GGC TGG CAG GAC TGG ATC ATC GCG 174
943 Arg Asp Leu Gly Trp Gln Asp Trp Ile Ile Ala
944 50 55
945 CCT GAA GGC TAC GCC GCC TAC TAC TGT GAG GGG 207
946 Pro Glu Gly Tyr Ala Ala Tyr Tyr Cys Glu Gly
947 60 65
948 GAG TGT GCC TTC CCT CTG AAC TCC TAC ATG AAC 240
949 Glu Cys Ala Phe Pro Leu Asn Ser Tyr Met Asn
950 70 75 80
951 GCC ACC AAC CAC GCC ATC GTG CAG ACG CTG GTC 273
952 Ala Thr Asn His Ala Ile Val Gln Thr Leu Val
953 85 90
954

Patent Application US/07/599,543B

955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986 CAC TTC ATC AAC CCG GAA ACG GTG CCC AAG CCC 306
987 His Phe Ile Asn Pro Glu Thr Val Pro Lys Pro
988 95 100
989 TGC TGT GCG CCC ACG CAG CTC AAT GCC ATC TCC 339
990 Cys Cys Ala Pro Thr Gln Leu Asn Ala Ile Ser
991 105 110
992 GTC CTC TAC TTC GAT GAC AGC TCC AAC GTC ATC 372
993 Val Leu Tyr Phe Asp Asp Ser Ser Asn Val Ile
994 115 120
995 CTG AAG AAA TAC AGA AAC ATG GTG GTC CGG GCC 405
996 Leu Lys Lys Tyr Arg Asn Met Val Val Arg Ala
997 125 130 135
998 TGT GGC TGC CAC TAGCTCCTCC GAGAATTCAG 437
999 Cys Gly Cys His

1000

1001

1002 (2) INFORMATION FOR SEQ ID NO:7:

1003 (i) SEQUENCE CHARACTERISTICS:

1004 (A) LENGTH: 102 amino acids

1005 (B) TYPE: amino acid

1006 (D) TOPOLOGY: linear

1007 (ii) MOLECULE TYPE: protein

Patent Application US/07/599,543B

1008 (ix) FEATURE:
1009 (D) OTHER INFORMATION:
1010 wherein each Xaa independently represents one of
1011 the 20 naturally occurring L-isomer, α -amino acids.
1012

1013 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:
1014
1015 Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
1016 1 5 10
1017 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
1018 15 20
1019 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa
1020 25 30
1021 Cys Xaa Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa
1022 35 40
1023 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
1024 45 50 55
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055

1056 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys
1057 60 65
1058 Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
1059 70 75
1060 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa

Patent Application US/07/599,543B

1061 80 85
1062 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys
1063 90 95
1064 Xaa Cys Xaa
1065 100
1066
1067 (2) INFORMATION FOR SEQ ID NO:8:
1068 (i) SEQUENCE CHARACTERISTICS:
1069 (A) LENGTH: 97 amino acids
1070 (B) TYPE: amino acid
1071 (D) TOPOLOGY: linear
1072 (ii) MOLECULE TYPE: protein
1073 (ix) FEATURE:
1074 (D) OTHER INFORMATION:
1075 wherein each Xaa independently represents one of
1076 the 20 naturally occurring L-isomer, α-amino acids
1077
1078 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:
1079
1080 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
1081 1 5 10
1082 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
1083 15 20
1084 Xaa Xaa Cys Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa
1085 25 30
1086 Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
1087 35 40
1088 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
1089 45 50 55
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113

Patent Application US/07/599,543B

1114
1115
1116
1117
1118
1119
1120
1121
1122 Xaa Xaa Xaa Xaa Xaa Cys Cys Xaa Xaa Xaa Xaa
1123 60 65
1124 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
1125 70 75
1126 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
1127 80 85
1128 Xaa Xaa Xaa Xaa Xaa Cys Xaa Cys Xaa
1129 90 95
1130
1131

- (2) INFORMATION FOR SEQ ID NO:9:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 136 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: protein
(ix) FEATURE:
(A) NAME: hOP-2P
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

1141
1142 Pro Leu Arg Arg Arg Gln
1143 1 5
1144 Pro Lys Lys Ser Asn Glu Leu Pro Gln
1145 10 15
1146 Ala Asn Arg Leu Pro Gly Ile Phe Asp
1147 20
1148 Asp Val Asn Gly Ser His Gly Arg Gln
1149 25 30
1150 Val Cys Arg Arg His Glu Leu Tyr Val
1151 35 40
1152 Ser Phe Gln Asp Leu Gly Trp Leu Asp
1153 45 50
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166

Patent Application US/07/599,543B

1220	Pro	Lys	Lys	Ser	Asn	Glu	Leu	Pro	Gln
1221		5					10		
1222									
1223									
1224									
1225									
1226									
1227									
1228									
1229									
1230									
1231									
1232									
1233									
1234									
1235									
1236									
1237									
1238									
1239									
1240									
1241									
1242									
1243									
1244									
1245									
1246									
1247									
1248									
1249									
1250									
1251									
1252									
1253	Ala	Asn	Arg	Leu	Pro	Gly	Ile	Phe	Asp
1254			15					20	
1255	Asp	Val	Asn	Gly	Ser	His	Gly	Arg	Gln
1256			25						30
1257	Val	Cys	Arg	Arg	His	Glu	Leu	Tyr	Val
1258			35						
1259	Ser	Phe	Gln	Asp	Leu	Gly	Trp	Leu	Asp
1260	40					45			
1261	Tyr	Val	Ile	Ala	Pro	Gln	Gly	Tyr	Ser
1262		50					55		
1263	Ala	Tyr	Tyr	Cys	Glu	Gly	Glu	Cys	Ser
1264			60					65	
1265	Phe	Pro	Leu	Asp	Ser	Cys	Met	Asn	Ala
1266			70						75
1267	Thr	Asn	His	Ala	Ile	Leu	Gln	Ser	Leu
1268					80				
1269	Val	His	Leu	Met	Lys	Pro	Asn	Ala	Val
1270	85					90			
1271	Pro	Lys	Ala	Cys	Cys	Ala	Pro	Thr	Lys
1272		95					100		

Patent Application US/07/599,543B

1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186 Tyr Val Ile Ala Pro Gln Gly Tyr Ser
1187 55 60
1188 Ala Tyr Tyr Cys Glu Gly Glu Cys Ser
1189 65
1190 Phe Pro Leu Asp Ser Cys Met Asn Ala
1191 70 75
1192 Thr Asn His Ala Ile Leu Gln Ser Leu
1193 80 85
1194 Val His Leu Met Lys Pro Asn Ala Val
1195 90 95
1196 Pro Lys Ala Cys Cys Ala Pro Thr Lys
1197 100 105
1198 Leu Ser Ala Thr Ser Val Leu Tyr Tyr
1199 110
1200 Asp Glu Ser Asn Asn Val Ile Leu Arg
1201 115 120
1202 Lys Ala Arg Asn Met Val Val Lys Ala
1203 125 130
1204 Cys Gly Cys His. *not valid, please delete*
1205 135
1206
1207

1208 (2) INFORMATION FOR SEQ ID NO:10:
1209 (i) SEQUENCE CHARACTERISTICS:
1210 (A) LENGTH: 133 amino acids
1211 (B) TYPE: amino acid
1212 (D) TOPOLOGY: linear
1213 (ii) MOLECULE TYPE: protein
1214 (ix) FEATURE:
1215 (A) NAME: hOP-2R
1216 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:
1217

1218 Arg Arg Gln
1219 1

Patent Application US/07/599,543B

1273 Leu Ser Ala Thr Ser Val Leu Tyr Tyr
1274 105 110
1275 Asp Glu Ser Asn Asn Val Ile Leu Arg
1276 115 120
1277 Lys Ala Arg Asn Met Val Val Lys Ala
1278 125
1279 Cys Gly Cys His *delete*
1280 130
1281
1282

1283 (2) INFORMATION FOR SEQ ID NO:11:

1284 (i) SEQUENCE CHARACTERISTICS:

1285 (A) LENGTH: 170 amino acids 160

1286 (B) TYPE: amino acid
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316

1317 (D) TOPOLOGY: linear

1318 (ii) MOLECULE TYPE: protein

1319 (ix) FEATURE:

1320 (A) NAME: hOP-2S

1321 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

1322
1323 Ser Gln Gln
1324 1
1325 Pro Phe Val Val Thr Phe Phe Arg Ala

Patent Application US/07/599,543B

1379
1380
1381
1382
1383
1384 Leu Ser Ala Thr Ser Val Leu Tyr Tyr
1385 130 135
1386 Asp Glu Ser Asn Asn Val Ile Leu Arg
1387 140 145
1388 Lys Ala Arg Asn Met Val Val Lys Ala
1389 150 165
1390 Cys Gly Cys His 155
1391 170
1392 160
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431

Page: 28

Raw Sequence Listing
Patent Application US/07/599,543B

07/08/91
11:20:47

1432
1433
1434

LINE ERROR

ORIGINAL TEXT

27 Wrong application Serial Number
538 Entered and Calc. Seq. Length differ *see p.15*
811 Response Exceeds Line Limitations
812 Response Exceeds Line Limitations
813 Response Exceeds Line Limitations
814 Response Exceeds Line Limitations
815 Response Exceeds Line Limitations
816 Response Exceeds Line Limitations
817 Response Exceeds Line Limitations
818 Response Exceeds Line Limitations
819 Response Exceeds Line Limitations
820 Response Exceeds Line Limitations
821 Response Exceeds Line Limitations
822 Response Exceeds Line Limitations
823 Response Exceeds Line Limitations
824 Response Exceeds Line Limitations
1204 Wrong Amino Acid Designator
1140 Entered and Calc. Seq. Length differ *error due to +0*
1279 Wrong Amino Acid Designator
1216 Entered and Calc. Seq. Length differ *error due to +0*
1321 Entered and Calc. Seq. Length differ

4 line limit, but o.k. here

OK → (A) APPLICATION NUMBER: US 07/599,543
→ (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:
at res. 19 = (Ile or Val); Xaa at res.23
Xaa at res.26 = (Ala or Ser); Xaa at res
or Ser); Xaa at res.38= (Asn or Asp); Xa
(Tyr or Cys); Xaa at res.49 = (Val or Le
res.52= (His or Asn); Xaa at res. 53 = (
Leu); Xaa at res. 54 = (Ile or Met); Xaa
or Lys); Xaa at res. 56 = (Glu, Asp or A
57=(Thr, Ala or Val); Xaa at res. 61 = (
Xaa at res. 67=(gln or Lys); Xaa at res.
(Asn or Ser); Xaa at 71=(Ile or Thr); Xa
76= (Phe or Tyr); Xaa at res. 78 = (Asp,
Xaa at res. 80= (Ser or Asn); Xaa at res
Asp); Xaa at res. 85 Arg); Xaa at res. 8
or His); and Xaa at res. 93=(Arg or Lys)
Cys Gly Cys His *delete*
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:
Cys Gly Cys His *0*
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

PAGE: 1

SEQUENCE MISSING ITEM REPORT
PATENT APPLICATION US/07/599,543B

DATE: 07/08/91
TIME: 11:20:48

MANDATORY IDENTIFIER THAT WAS NOT FOUND

CLASSIFICATION

LINE ORIGINAL TEXT

CORRECTED TEXT

78 (A) NAME: mOP2 (mature)
157 (A) NAME: mOP2
425 (A) NAME: hOP2 (mature)
537 (A) NAME: hOP2
924 (A) NAME: OP1
1139 (A) NAME: hOP-2P
1215 (A) NAME: hOP-2R
1320 (A) NAME: hOP-2S

(A) NAME/KEY: mOP2 (mature)
(A) NAME/KEY: mOP2
(A) NAME/KEY: hOP2 (mature)
(A) NAME/KEY: hOP2
(A) NAME/KEY: OP1
(A) NAME/KEY: hOP-2P
(A) NAME/KEY: hOP-2R
(A) NAME/KEY: hOP-2S